Soil Biology & Biochemistry 76 (2014) 53–56

Contents lists available at ScienceDirect

Soil Biology & Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Short communication

High-throughput sequencing shows inconsistent results with a microscope-based analysis of the soil prokaryotic community

Masayuki Ushio [a,*], Kobayashi Makoto [b,c,d], Jonatan Klaminder [b], Hiroyuki Takasu [a,e], Shin-ichi Nakano [a]

[a] Center for Ecological Research, Kyoto University, 2-509-3 Hirano, Otsu, Shiga 520-2113, Japan
[b] Climate Impact Research Centre, Umeå University, 981-07, Abisko, Sweden
[c] Soil Ecology Research Group, Graduate School of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Yokohama, Kanagawa 240-8501, Japan
[d] Nakagawa Experimental Forest, Field Science Center for Northern Biosphere, Hokkaido University, Otoineppu 098-2501, Japan
[e] Atmosphere and Ocean Research Institute (AORI), The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan

ARTICLE INFO

Article history:
Received 7 February 2014
Received in revised form
1 May 2014
Accepted 3 May 2014
Available online 19 May 2014

Keywords:
Archaea
Bacteria
CARD–FISH
High-throughput sequencing
Soil prokaryotic community
Tundra ecosystem

ABSTRACT

In the present study, we perform the first direct analysis on how the composition of the prokaryotic soil community differs depending on whether high-throughput sequencing or fluorescent *in situ* hybridization (FISH) coupled with catalyzed reporter deposition (CARD) is used. Soil samples were collected along short (<3 m) tundra vegetation gradients from Northern Sweden. Relative abundances of *Acidobacteria* and *Bacteroidetes* estimated by the high-throughput sequencing were higher than those estimated by CARD–FISH, while relative abundances of *Archaea* and *α-Proteobacteria* estimated by high-throughput sequencing were lower than those estimated by CARD–FISH. The results indicated that the high-throughput sequencing overestimates/underestimates the relative abundance of some microbial taxa if we assume that CARD–FISH can provide potentially more quantitative data. Great caution should be taken when interpreting data generated by molecular technologies (both of high-throughput sequencing and CARD–FISH), and supports by multiple approaches are necessary to make a robust conclusion.

When investigating the composition of microbial communities, high-throughput sequencing technologies are becoming the most commonly applied method in microbial ecology. The rationale for this is a cost-effective means of identifying thousands of microbial phylotypes that are present in samples (Sogin et al., 2006; Lauber et al., 2009). Without these technologies, it is almost impossible to reveal the very high diversity of soil microbial communities, and thus, they currently constitute the most important tools for our understanding about soil microbes.

However, many experimental steps in the high-throughput sequencing analysis could potentially produce biases/artifacts that significantly influence biological interpretations of the dataset (Gomez-Alvarez et al., 2009; Engelbrektson et al., 2010; Zhou et al., 2011). For example, Zhou et al. (2011) examined the quantitative capacity of high-throughput sequencing of 16S rRNA gene amplicons by adding a known quantity of extracted DNA from a cultured microbial strain. They found that the percentages of the strain OTU varied substantially among different samples, from 0.00% to 5.34% (theoretically it should have been 0.1%), and thus they questioned the quantitative capacity of high-throughput sequencing. However, studies that compared results of other potentially more quantitative approaches (e.g., microscope-based investigations) in complex soil matrixes have been very limited so far.

To compare the results of high-throughput sequencing and a potentially more quantitative microscope-based analysis, we analyzed soil samples with IonPGM high-throughput sequencer (Rothberg et al., 2011; Ion Torrent by Life Technologies, Guilford, CT, USA) and fluorescent *in situ* hybridization (FISH) coupled with catalyzed reporter deposition (CARD) (Pernthaler et al., 2002; Eickhorst and Tippkötter, 2008), targeting *Bacteria* and *Archaea*. In the present study, we focused on the methodological comparison because the results of CARD–FISH analysis of the soil prokaryotic community were already reported in Ushio et al. (2013). Soil samples were collected from the north-facing slope of Mt.

* Corresponding author. Tel.: +81 77 549 8215; fax: +81 77 549 8201.
E-mail addresses: ong8181@gmail.com, ushima@ecology.kyoto-u.ac.jp (M. Ushio).

http://dx.doi.org/10.1016/j.soilbio.2014.05.010
0038-0717/

Suorooaivi in Abisko, Northern Sweden. In this area, patterned ground also referred to as non-sorted circles (Fig. 1a; Klaus et al., 2013) occurs frequently because of the soil-frost process. Within these features, a dramatic vegetation change from lichen-dominated plant communities to dense shrub communities occurs over distances of less than 3 m (Fig. 1b; Makoto and Klaminder, 2012).

Soil samples were collected along the vegetation gradients (6 individual circles × 7 locations = 42 samples; Fig. 1b) and taken back to the laboratory immediately. After the sorting, CARD–FISH was conducted as described previously (Ushio et al., 2013; Supplementary methods), and the soil subsamples were stored at −20 °C until further DNA analyses. For the CARD–FISH analysis, nine probes (*Eubacteria*, *Archaea*, *α-Proteobacteria*, *β-Proteobacteria*, *γ-Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *CFB*, and Nonsense probes), which target potentially dominant soil microbial groups (e.g., Lauber et al., 2009), were applied to quantify the abundance of microbial groups in the soil samples (see Table 1 in Ushio et al., 2013). By following Bates et al. (2011), soil DNA extraction, polymerase chain reaction (PCR), and purification were conducted (see detail in Supplementary methods). The multiple PCR products were pooled, and the single composite 6-bp-barcoded sample was sent for sequencing at Life Technologies Japan (Tokyo, Japan) on IonPGM. After sequencing, the raw sequence data were processed using QIIME (Caporaso et al., 2010). Quality filtering, chimera identifications, and operational taxonomic unit (OTU) clustering (≥97% similarity) were performed using the USEARCH option (Edgar, 2010; Edgar et al., 2011) in QIIME. After filtering and clustering, taxonomies were assigned to the OTUs. Detailed experimental protocols and data handling procedures are described in the Supplementary information. The sequences data is archived in DNA Data Bank of Japan (DDBJ) Sequence Read Archives (DRA), and the accession number is DRA001218 for the submission data.

After processing, 620,345 sequences from a total of 42 soil samples were obtained (14,770 ± 4564 [±standard deviation] sequences per sample; Table S1), and approximately 2200 OTUs were identified for each sample (Table S1 and Fig. S1). According to the high-throughput sequencing analysis, *Acidobacteria* was the most dominant microbial group, followed by *Bacteroidetes*, *α-Proteobacteria*, and *γ-Proteobacteria* (Fig. 1c). The relative abundance of *Archaea* was less than 1% (Fig. 1c). Conversely, the relative abundance of *Archaea* estimated by CARD–FISH exceeded 30% in the soil samples (Fig. 1d), which is similar in range to that previously reported for farmland and paddy soils by CARD–FISH (Eickhorst and

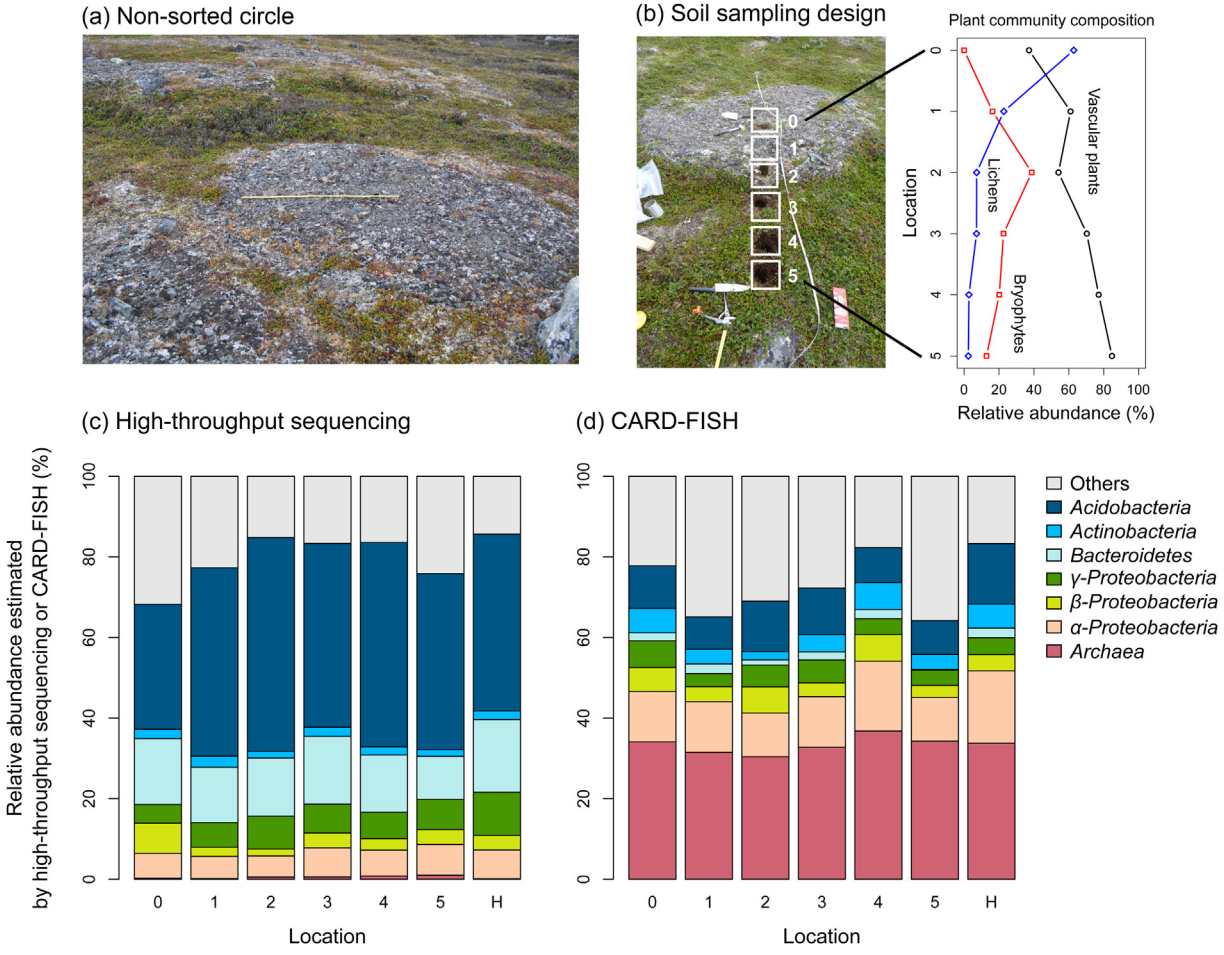

Fig. 1. (a) Images of a non-sorted circle (NSC), where a 1-m measure is placed between the center of the NSC and the edge of the inner domain. (b) Soil sampling design (left panel). Here, six sampling points were assigned between the center of NSC and the edge of the outer domain. Aboveground vegetation is absent or relatively poor at locations 0, 1, and 2, while plants are densely colonized at locations 3, 4, and 5. The humus layer, which is present only at locations 3, 4, and 5, was also collected. This sampling design is identical to that of Makoto and Klaminder (2012). Plant community composition changed dramatically along this transect because of the soil-frost process (cryoturbation) in the system (right panel). (c) Composition of the prokaryotic community along the NSC vegetation gradients estimated by high-throughput sequencing, or (d) CARD–FISH. Numbers on the *x*-axis indicate distance from the center of the NSC, and "H" indicates humus layer samples. Each bar represents the mean value of the relative abundance of each microbial group. "Others" includes prokaryotic microbes other than the listed microbial groups, and unidentified sequence reads. (a), (b), and (d) are reproduced and modified from Ushio et al. (2013) with the permission from Elsevier.